CHROM. 5929

A STUDY OF THE CHROMATOGRAPHIC PROPERTIES OF DIPEPTIDES BY AUTOMATIC ION-EXCHANGE CHROMATOGRAPHY

C. HAWORTH
Department of Chemistry and Applied Chemistry, The University, Salford 5 Lancs. (Great Britain)

(Received January 20th, 1972)

SUMMARY

A new empirical relationship for calculating the relative elution volumes of dipeptides from those of the constituent amino acids is proposed. The relationship is based on the additivity of relative elution volumes.

It is shown that values determined using the new equation agree closely with the observed and reported elution volumes. The reason for small discrepancies, occurring in a few cases, between the observed values and calculated values are discussed.

INTRODUCTION

The identification of low molecular weight peptides is of importance in the field of biochemistry. For example, it is an essential step in the determination of the structure of a protein from the protein hydrolysate. In body fluids the occurrence of low molecular weight peptides is of clinical interest because their presence may be correlated with renal physiology and metabolism[1]. However, the techniques required for the identification of even di- and tripeptides occurring in such complex mixtures are time-consuming and present many varied problems. The general approach to solving the problems is initially to obtain a 'map' of the peptide containing solution using paper or thin-layer chromatographic techniques. Each individual spot on the map is then separately eluted for complete isolation.

In order to obtain a pure compound for characterisation each individual spot is subjected to further separation techniques using either paper (PC), thin-layer (TLC), or automatic ion-exchange chromatography. Because of the popularity of these techniques attempts have been made to correlate the PC and TLC properties of low molecular weight peptides with those of the constituent amino acids[2–5]. An empirical equation relating such chromatographic properties on paper and thin layers of cellulose has been obtained by HAWORTH AND OLIVER[5] which gave results generally in excellent agreement with the observed values. This empirical equation has been used in the identification of carnosine in the urine of a mentally retarded

child[6]. Despite the work on paper and thin layers no comparable study using an automatic ion-exchange technique has been reported. Recently a graphical interpretation has been made by means of which it was possible to predict the elution position of some dipeptides[2]. Accordingly it was decided to see if a general relationship between the chromatographic properties of these peptides and their constituent amino acids exists for the automatic ion-exchange technique. As a starting point in this study the properties of dipeptides on the Technicon instrument were investigated. The results obtained are presented in this paper and show that a simple empirical relationship between the dipeptides and the constituent amino acids does exist.

EXPERIMENTAL

Materials and equipment

Automatic ion exchange. The instrument used was the Technicon* auto amino acid analyser Model NC-1. The column (150 × 0.9 cm) was filled with the sulphonated polystyrene resin Chromobeads (Type B) which has a degree of cross-linking equivalent to 8% divinyl benzene (D.V.B.) content.

Amino acids and peptides. With the exception of DL-alanyl-DL-alanine, DL-leucyl-DL-valine, and DL-alanyl-DL-serine, the amino acids and dipeptides used in this work were the L-enantiomorphs and obtained from a variety of commercial sources.

Standard solutions. Stock solutions (0.025 *M*) of the amino acids and dipeptides were made up in 2-propanol in water (10%). The amount of each dipeptide applied to the auto analyser necessary to show an identifiable peak had to be determined individually by control experiments[7].

Methods

Operation of Technicon auto amino acid analyser. The 24 h chromatogram employing sodium citrate buffers at a pH gradient 2.875–5.000 and a flow rate of 30 ml/h was used according to the instructions given in the Technicon operation manual[8]. The amino acids and dipeptide solutions were applied to the column together with an equal volume of pH 2.875 citrate buffer.

Since arginine is the last of the commonly occurring amino acids to emerge from the column it was chosen as arbitrary standard throughout this work. The relative retention volume of arginine (R_{Arg}) was given the value 100 and the elution volumes of other substances were measured relative to arginine. For the purpose of simplification R_{Arg} is abbreviated to R throughout this work.

RESULTS AND DISCUSSION

Table I gives the R values of eight amino acids and ten dipeptides obtained during the course of this work, together with the R values of seventeen amino acids and twenty-three dipeptides converted from the results given by other authors[7].

* Technicon Instruments Co. Ltd., Hamilton Close, Houndsmills, Basingstoke, Hants.

TABLE I

R VALUES[a] OF TWENTY-FIVE AMINO ACIDS AND THIRTY-THREE SYNTHETIC DIPEPTIDES AFTER CHROMATOGRAPHY USING THE TECHNICON AUTO ANALYSER (24 h Chromatogram)

No.	*Compound*	*R value*	*No.*	*Compound*	*R value*
1	Alanine	31.4	30	Alanyl–phenylalanine	75.2
2	Aspartic acid	13.5	31	Alanyl–serine	34.0
3	Asparagine	18.1	32	β-Alanyl–histidine	86.7
4	Arginine	100.0	33	Glutamyl–alanine	39.8
5	β-Alanine	61.9	34	Glycyl–alanine	47.9
6	Cystine	45.1	35	Glycyl–aspartic acid	33.9
7	Cysteic acid	3.6	36	Glycyl–glycine	35.2
8	Glutamic acid	21.0	37	Glycyl–isoleucine	64.3
9	Glutamine	18.1	38	Glycyl–leucine	64.8
10	Glycine	28.2	39	Glycyl–lysine	94.3
11	Histidine	84.4	40	Glycyl–proline	57.1
12	Hydroxyproline	9.6	41	Glycyl–serine	29.3
13	Isoleucine	52.0	42	Glycyl–tyrosine	73.2
14	Leucine	54.3	43	Glycyl–valine	55.3
15	Lysine	80.9	44	Leucyl–alanine	55.5
16	Methionine	46.6	45	Leucyl–glycine	63.3
17	1-Methylhistidine	79.5	46	Leucyl–valine	60.7
18	3-Methylhistidine	84.3	47	Seryl–glycine	38.5
19	Phenylalanine	60.6	48	Valyl–alanine	52.7
20	Proline	23.2	49	Valyl–glycine	51.9
21	Serine	16.4	50	Valyl–histidine	93.0
22	Threonine	15.8	51	Valyl–leucine	65.9
23	Tryptophan	88.7	52	Valyl–methionine	64.8
24	Tyrosine	58.6	53	Valyl–phenylalanine	77.4
25	Valine	41.1	54	Valyl–proline	60.2
26	Alanyl–alanine	47.5	55	Valyl–serine	45.7
27	Alanyl–aspartic acid	31.6	56	Valyl–tryptophan	99.3
28	Alanyl–glutamic acid	42.3	57	Valyl–tyrosine	74.5
29	Alanyl–glycine	44.7	58	Valyl–valine	61.0

[a] *R* = elution volume relative to arginine.

The *R* values have been converted from R_{AH} values[7,*] using the observed elution volumes of aspartic acid (170 ml) and histidine (1065 ml). Several of the reported values have been checked during this work and found to be in good agreement.

It has previously been shown that for TLC a plot of the R_F value of the dipeptide against the R_F value of the C-terminal amino acid gave a linear relationship for a constant N-terminal amino acid. Consequently a plot of the *R* value for the constant N-terminal series of glycyl, alanyl and valyl dipeptides against the *R* values of the constituent C-terminal amino acids was made for automatic ion exchange. The plots obtained are shown in Fig. 1 as full lines. Since it has been shown that structural isomers have very similar properties on paper, thin layers and ion-exchange chromatographic systems in most cases[2,4,5,7] plots can be made for seryl, tyrosyl, phenylalanyl and histidyl dipeptides and these are shown in Fig. 1 as dotted lines. From a study of Fig. 1 it can be seen that in each case straight line plots were obtained for dipeptides

* $R_{AH} = \frac{\text{elution time (min) of peak}}{\text{elution time (min) between standards}}$.

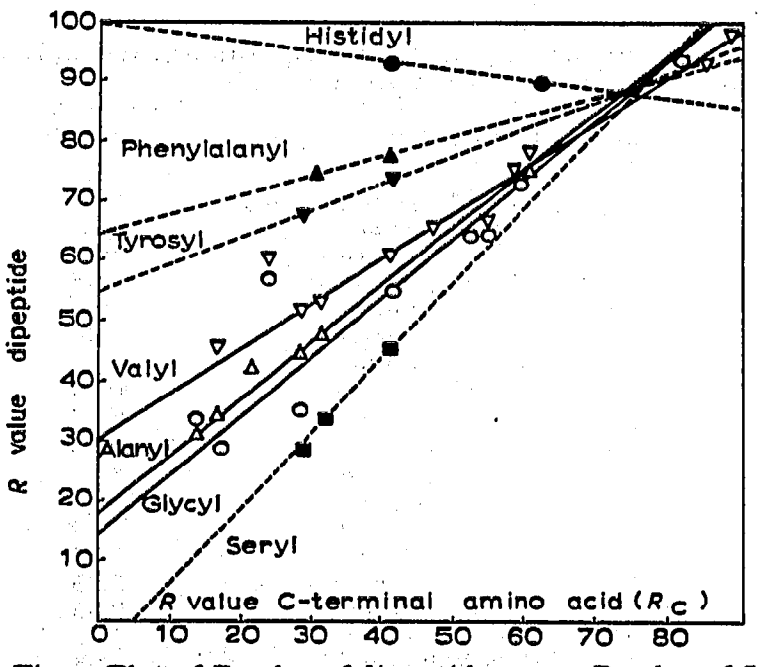

Fig. 1. Plot of R value of dipeptide *versus* R value of C-terminal amino acid for the Technicon auto analyser. Symbols: ○, glycyl; △, alanyl; ▽, valyl; ●, histidyl; ▼, tyrosyl; ▲, phenylalanyl; ■, seryl; - - - - -, plots obtained assuming that structural isomers have similar chromatographic properties.

having constant N-terminal amino acid residues. Thus for a particular series from the equation of a straight line

$$R_{Pep} = mR_C + I \qquad (1)$$

where R_{Pep} is the R value of the dipeptide, R_C is the R value of the constituent C-terminal amino acid, m is the gradient and I is the intercept on the abscissa. Because the plots obtained (Fig. 1) are straight lines m and I are not functions of R_C. Neither m nor I were constant from one series to another, and thus both were found to vary with the nature of the N-terminal amino acid. Therefore m and I are functions of the R value of the N-terminal amino acid (R_N). Consequently it may be written:

$$m \approx F(R_N)$$
$$\text{and } I \approx F(R_N)$$

From Fig. 1, m and I have been calculated and are given in Table II for each series studied together with the R values of the corresponding free amino acid. A study of Table II shows that the gradient m decreases with an increase in the R value of the N-terminal amino acid. The converse applies in the case of the intercept I which increases with an increase in the R value of the N-terminal amino acid. In an attempt to evaluate the above functions a plot of m *versus* R_N and a plot of I *versus* R_N were made using the data given in Table II. The plots obtained are given in Figs. 2 and 3 respectively. From Fig. 2, the straight line obtained shows that the gradient m is directly proportional to R_N. Therefore

$$m = m' R_N + I' \qquad (2)$$

where m' was found to be −0.0191 and I' was found to be 1.50.

TABLE II

GRADIENT (m) AND INTERCEPT (I) ON ABSCISSA OBTAINED FROM STRAIGHT LINE PLOTS OF R VALUE OF THE DIPEPTIDE *versus* THE R VALUE OF THE C-TERMINAL AMINO ACID (Fig. 1)

N-Terminal amino acid	*Gradient (m)*	*Intercept (I)*	*R value*
Serine[a]	+1.25	−5.5	16.4
Glycine	+0.99	15.0	28.2
Alanine	+0.94	18.5	31.4
Valine	+0.75	30.5	41.1
Tyrosine[a]	+0.45	55.0	58.6
Phenylalanine[a]	+0.33	65.0	60.6
Histidine[a]	−0.17	100.0	84.4

[a] Gradient and intercept values from plots obtained assuming that structural isomers elute at the same position.

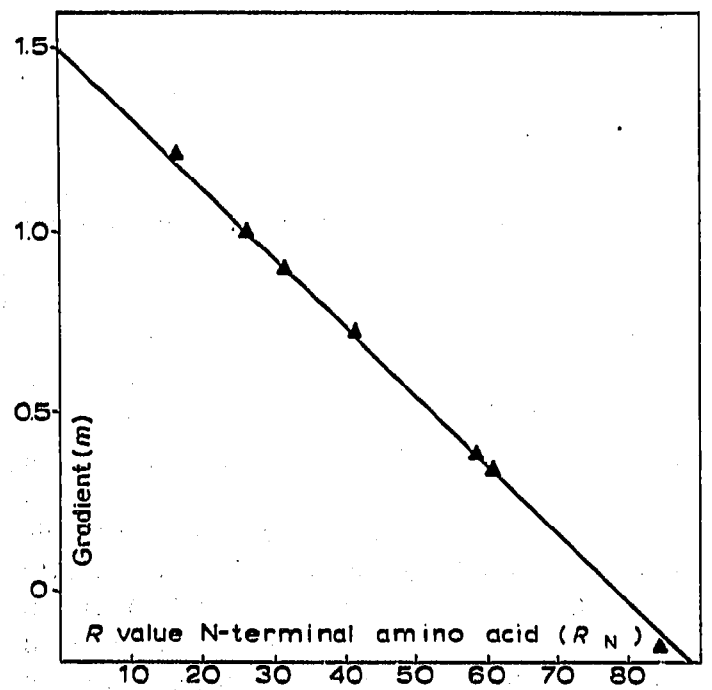

Fig. 2. Plot of gradient (m) *versus* the R value of the N-terminal amino acid for the series of N-terminal amino acids given in Table II.

Similarly the straight line obtained in Fig. 3 shows that the intercept I is also directly proportional to R_N. Therefore

$$I = m'' R_N + I'' \quad (3)$$

where m'' was found to be 1.50 and I'' was found to be −28.5.

Substituting eqns. 2 and 3 in eqn. 1 the following is obtained

$$R_{Pep} = R_C[m' R_N + I'] + [m'' R_N + I''] \quad (4)$$

Inserting the values for m', I', m'' and I'' calculated from Figs. 2 and 3 the equation becomes

$$R_{Pep} = R_C[-0.0191\, R_N + 1.50] + [1.50\, R_N - 28.5]$$

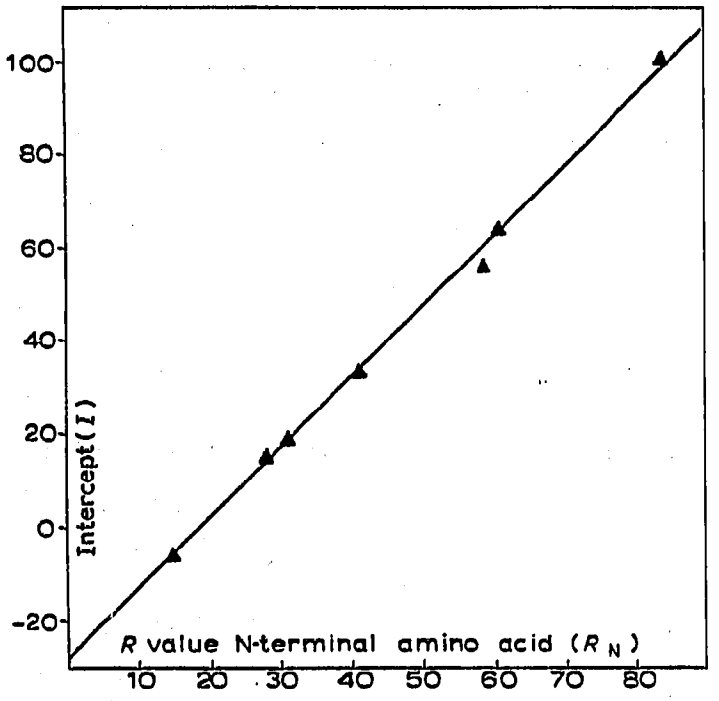

Fig. 3. Plot of intercept (I) *versus* the R value of the N-terminal amino acid for the series of N-terminal amino acids given in Table II.

This rearranges to

$$R_{Pep} = 1.50[R_N + R_C] - [0.0191\, R_N R_C] - 28.5 \quad (5)$$

This is the general equation relating the R_{Pep} of any dipeptide with the R values of the constituent amino acids.

Using this equation (5) the R_{Pep} values were recalculated for the available dipeptides and are compared with the observed values in Table III. From a study of this table it can be seen that the calculated R value was within ± 3 R units in twenty-one out of thirty-three cases and within ± 5 R units in twenty-five out of thirty-three cases. The reproducibility of duplicate determinations of dipeptides on the Technicon instrument is not known and therefore I have chosen ± 5 R units as the criterion for correctness of the equation. The exceptional value outside these limits may be due to several factors which are not allowed for in the general equation. For example, the assumption that structural isomers possessed similar chromatographic properties is clearly incorrect in the case of the glycylserine and serylglycine pairs. The R value of glycylserine is closely predicted but that of serylglycine is not. This discrepancy may be explained as follows. Serine is structurally related to alanine but because of the hydroxyl group on the β-carbon atom it is of more acidic character than alanine. Consequently the R value of serine on the Technicon auto analyser is lower than that of alanine (see Table I) by ΔR of 15.0 units. However comparison of the R value of serylglycine and alanylglycine shows a ΔR of only 6.2 units. It is therefore apparent that the seryl group in the peptide is less acidic than is the free amino acid. This is probably due to intramolecular interaction with the peptide bond of substituent groups. For example, hydrogen bonding can occur between the hydroxyl group of the serine and the carbonyl group of the peptide bond (see Fig. 4).

TABLE III

COMPARISON OF OBSERVED R VALUES AND R VALUES CALCULATED USING THE EMPIRICAL EQN. 5 FOR THIRTY-THREE SYNTHETIC DIPEPTIDES AFTER CHROMATOGRAPHY USING THE TECHNICON AUTO ANALYSER

Dipeptide	*R value*		*Obs–calc.*
	Obs.	*Calc.*	*Δ*
Alanyl–alanine	47.5	46.9	+ 0.6
Alanyl–aspartic acid	31.6	30.8	+ 0.8
Alanyl–glutamic acid	42.3	37.5	+ 4.8
Alanyl–glycine	44.7	44.0	+ 0.7
Alanyl–phenylalanine	75.2	73.2	+ 2.0
Alanyl–serine	34.0	33.4	+ 0.6
β-Alanyl–histidine	86.7	91.5	− 4.8
Glutamyl–alanine	39.8	37.5	+ 2.3
Glycyl–alanine	47.9	44.0	+ 3.9
Glycyl–aspartic acid	33.9	26.7	+ 7.2
Glycyl–glycine	35.2	40.8	− 5.6
Glycyl–isoleucine	64.3	63.8	+ 0.5
Glycyl–leucine	64.8	66.0	− 1.2
Glycyl–lysine	94.3	91.5	+ 2.8
Glycyl–proline	57.1	21.4	+35.7
Glycyl–serine	29.3	29.6	− 0.3
Glycyl–tyrosine	73.2	70.1	+ 3.1
Glycyl–valine	55.3	53.3	+ 2.0
Leucyl–alanine	55.5	67.4	−11.9
Leucyl–glycine	63.3	66.0	− 2.7
Leucyl–valine	60.7	72.0	−11.3
Seryl–glycine	38.5	29.6	+ 8.9
Valyl–alanine	52.7	55.6	− 2.9
Valyl–glycine	51.9	53.3	− 1.4
Valyl–histidine	93.0	93.4	− 0.4
Valyl–leucine	65.9	72.0	− 6.1
Valyl–methionine	64.8	66.4	− 1.6
Valyl–phenylalanine	77.4	76.4	+ 1.0
Valyl–proline	60.2	50.0	+10.2
Valyl–serine	45.7	44.8	+ 0.9
Valyl–tryptophan	99.3	96.9	+ 2.4
Valyl–tyrosine	74.5	75.0	− 0.5
Valyl–valine	61.0	62.6	− 1.6

Also from Fig. 4 it can be seen that intramolecular action with the normal peptide bond cannot easily occur with glycylserine and consequently the observed and calculated values are found to agree well.

Reasons similar to the above may be proposed for the other dipeptides with

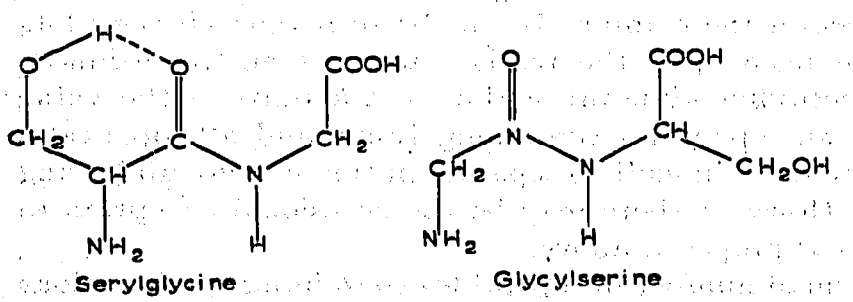

Fig. 4. Diagrammatic representation for structural isomers of dipeptides from the amino acids serine and glycine.

chromatographic properties not closely predicted using the equation, *e.g.*, the different type of peptide bond in those dipeptides containing C-terminal proline. However insufficient data is available to enable these effects to be more closely evaluated.

If the above factors are taken into account then eqn. 5 is valid and can be used to calculate the elution values of any dipeptides on the Technicon instrument.

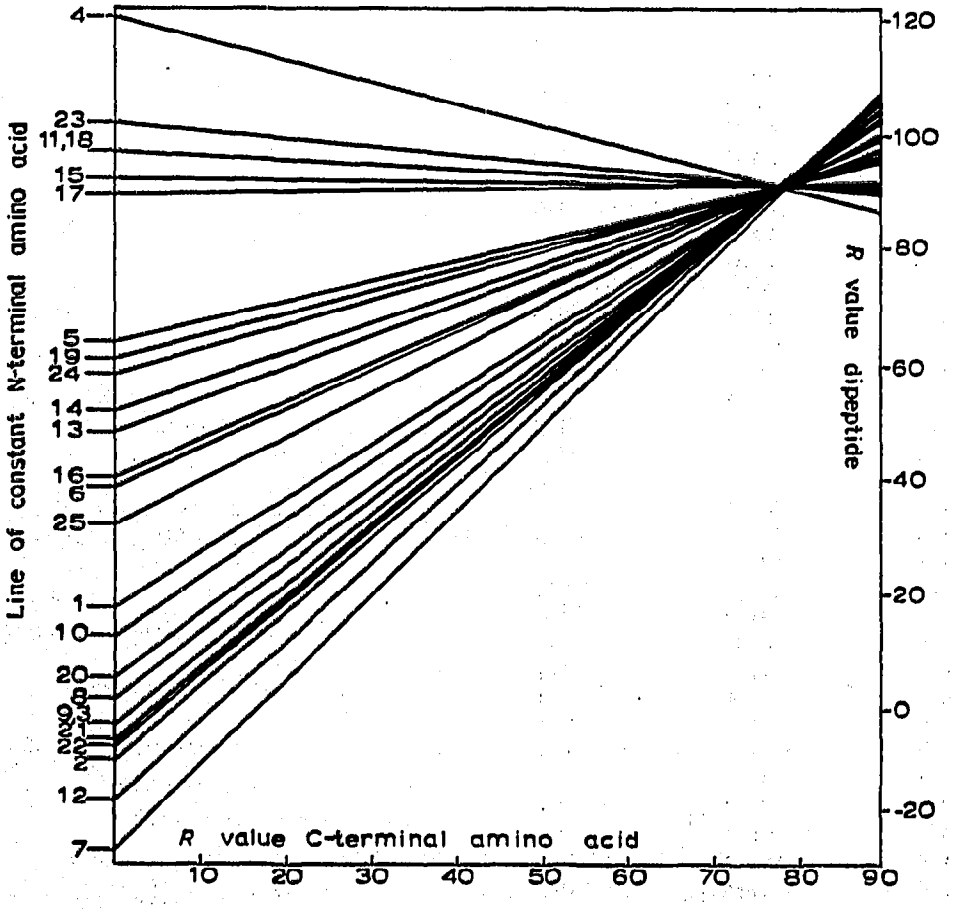

Fig. 5. Plot of R value of the dipeptide *versus* R value of the C-terminal amino acid for twenty-four of the more commonly occurring amino acids, calculated using eqn. 5. For key to the numbering system, see Table I.

Fig. 5 shows the plots of R value of the dipeptide *versus* the R value of the C-terminal amino acid for twenty-four more commonly occurring amino acids calculated using eqn. 5. The closer the plots are together in Fig. 5 the more difficult is the predicted separation of the peptides on the column. It can be seen that all the plots intersect at a point with coordinates 89.3 on the abscissa and 78.5 on the ordinate. The R values of lysine and 1-methylhistidine are within ± 5 R units of the value 78.5. This is significant in that all dipeptides containing lysine and all dipeptides containing 1-methylhistidine should be difficult to separate in this system, all having R values of about 89.3 R units. However there may be the occasional exception to this due to intramolecular forces as proposed above.

Also it can be seen that a small number of dipeptides containing combinations of the more acidic amino acids have R values calculated to be negative. Such dipeptides should be taken as eluting close to the void volume of the column.

It is concluded that a knowledge of the R value of a peptide could reduce the amount of time necessary to characterise it and this fact is of importance in sequence analysis of proteins and polypeptides. This is particularly the case if use of the ion-exchange eqn. 5 is combined with the use of a second separation technique, based on a different physico-chemical technique such as TLC for which a similar equation is available[5].

Such a combination of data should provide a reliable and powerful means for identification of unknown dipeptides.

ACKNOWLEDGEMENTS

I wish to thank Dr. J. G. HEATHCOTE for giving me a copy of the original manuscript of his report[7], Professor G. R. RAMAGE for the provision of laboratory facilities, and Mrs. M. HAWORTH for typing and proof reading this report.

REFERENCES

1 C. E. DENT, *Biochem. J.*, 43 (1948) 169.
2 P. X. CALLAHAN, J. A. SHEPARD, T. J. REILLY, K. MCDONALD AND S. ELLIS, *Anal. Biochem.*, 38 (1970) 380.
3 A. B. PARDEE, *J. Biol. Chem.*, 190 (1951) 757.
4 C. HAWORTH AND R. W. A. OLIVER, *Biochem. J.*, 124 (1971) 255.
5 C. HAWORTH AND R. W. A. OLIVER, *J. Chromatogr.*, 64 (1972) 305.
6 C. HAWORTH, *Ph.D. Thesis*, University of Salford, Great Britain, 1971.
7 J. G. HEATHCOTE, R. J. WASHINGTON, B. J. KEOGH AND R. W. GLANVILLE, *J. Chromatogr.*, 65 (1972) 397.
8 P. B. HAMILTON, in DORIS I. SCHMIDT (Editor), *Techniques in Amino Acid Analysis*, Technicon International Division, Geneva, 1966, p. 17.